

1 **TITLE**

2

3 Diversity of cytosine methylation across the fungal tree of life

4

5 **AUTHORS**

6

7 Adam J. Bewick¹, Brigitte T. Hofmeister², Rob Powers³, Stephen J. Mondo⁴, Igor
8 V. Grigoriev^{4,5}, Timothy Y. James³, Jason E. Stajich⁶, Robert J. Schmitz¹

9

10 **AFFILIATIONS**

11

12 ¹Department of Genetics, University of Georgia, Athens, Georgia, USA. ²Institute
13 of Bioinformatics, University of Georgia, Athens, Georgia, USA. ³Department of
14 Ecology and Evolutionary Biology, University of Michigan, Michigan, USA. ⁴US
15 Department of Energy Joint Genome Institute, Walnut Creek, California, USA.
16 ⁵Department of Plant and Microbial Biology, University of California, Berkeley,
17 Berkeley, California, USA. ⁶Department of Microbiology and Plant Pathology,
18 University of California, Riverside, Riverside, California, USA.

19

20 **CORRESPONDING AUTHORS**

21

22 Robert J. Schmitz schmitz@uga.edu and Adam J. Bewick bewickaj@uga.edu

23

24

25

26

27

28

29

30

31

32 **SUPPLEMENTARY INFORMATION**

33

34 **Supplementary Figure 1**

35

36 **5mC DNA and tRNA MTase phylogeny of fungi.**

37

38 Evolution of DNA and tRNA methyltransferases across fungi. Values at selected
39 nodes indicate posterior probability. Nodes with a star specify duplications, and
40 the single node with a diamond specifies the clade containing '*DnmtX*' (29).
41 Branch lengths are in units of amino acid substitutions per amino acid site. Phyla
42 abbreviations: Ascomycota (Asco-), Basidiomycota (Basidio-),
43 Blastocladiomycota (Blastocladio-), Chytridiomycota (Chytridio-), Mucoromycota
44 (Mucoro-), and Zoopagomycota (Zoopago-).

45

46

47

48

49

50

51

52

53

54

55

56

57

58

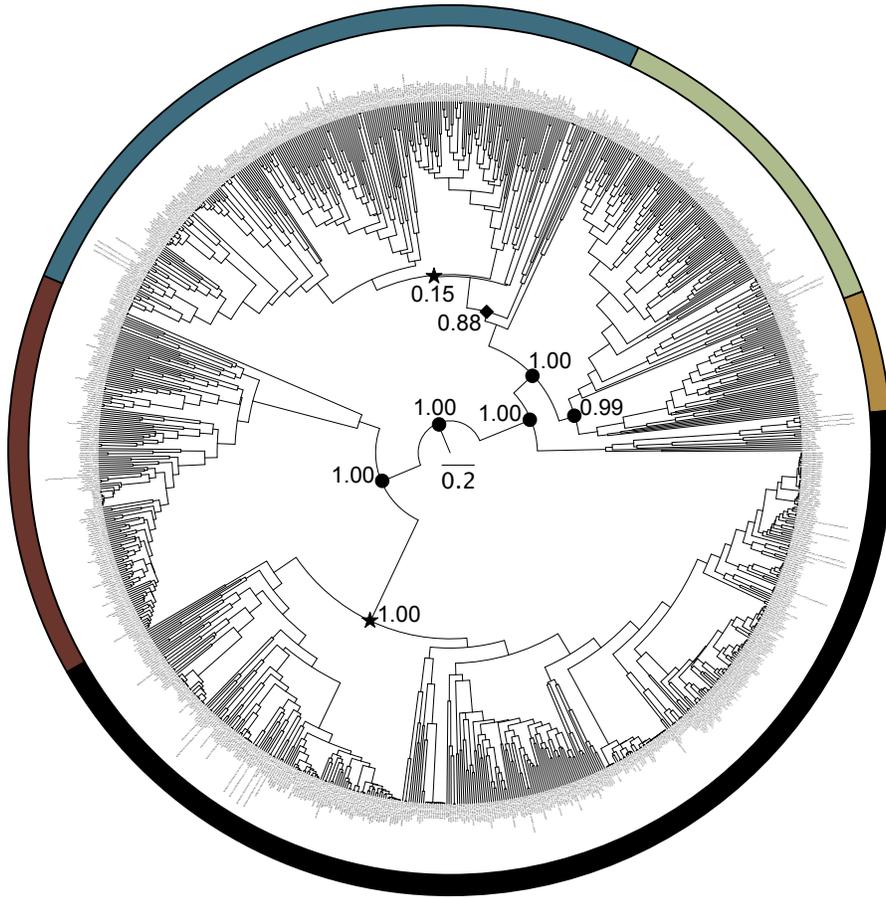
59

60

61

62

Supplementary Fig. 1



MTase

- | | | |
|---|---|--|
|  DNMT1 |  DNMT5 |  DIM-2 |
|  RID |  DNMT2 | |

63

64

65

66

67

68 **Supplementary Figure 2**

69

70 **5mC DNA and tRNA MTase phylogeny of animalia, chlorophyta, fungi, and**
71 **prokaryota.**

72

73 Evolution of DNA and tRNA methyltransferases across the tree of life. Values at
74 selected nodes indicate posterior probability. Branch lengths are in units of amino
75 acid substitutions per amino acid site. Phyla abbreviations: Ascomycota (Asco-),
76 Basidiomycota (Basidio-), Blastocladiomycota (Blastocladio-), Chytridiomycota
77 (Chytridio-), Mucoromycota (Mucoro-), and Zoopagomycota (Zoopago-).

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

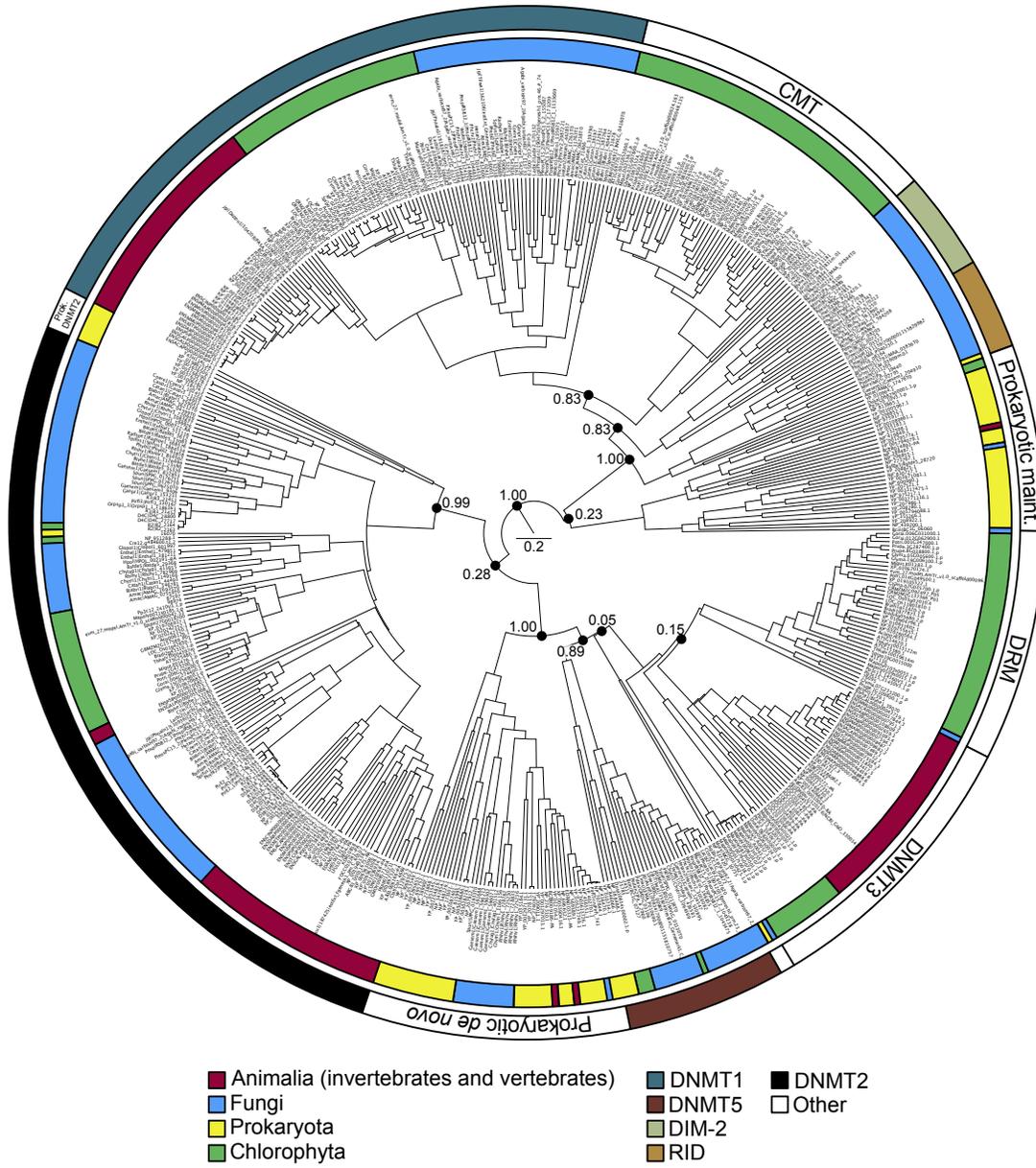
95

96

97

98

Supplementary Fig. 2



99
 100
 101
 102
 103

104 **Supplementary Figure 3**

105

106 **Stochastic character mapping of sixteen observed DNA MTase genotypes.**

107

108 The colored bars indicate phylum of encompassing fungal species. Phyla
109 abbreviations: Ascomycota (Asco-), Basidiomycota (Basidio-),
110 Blastocladiomycota (Blastocladio-), Chytridiomycota (Chytridio-), Mucoromycota
111 (Mucoro-), and Zoopagomycota (Zoopago-). The colored circle adjacent to each
112 species corresponds to the observed genotype. Pie charts at nodes indicate
113 empirical Bayesian posterior probabilities of the genotypes in a common
114 ancestor. Branch lengths are in units of amino acid substitutions per amino acid
115 site.

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

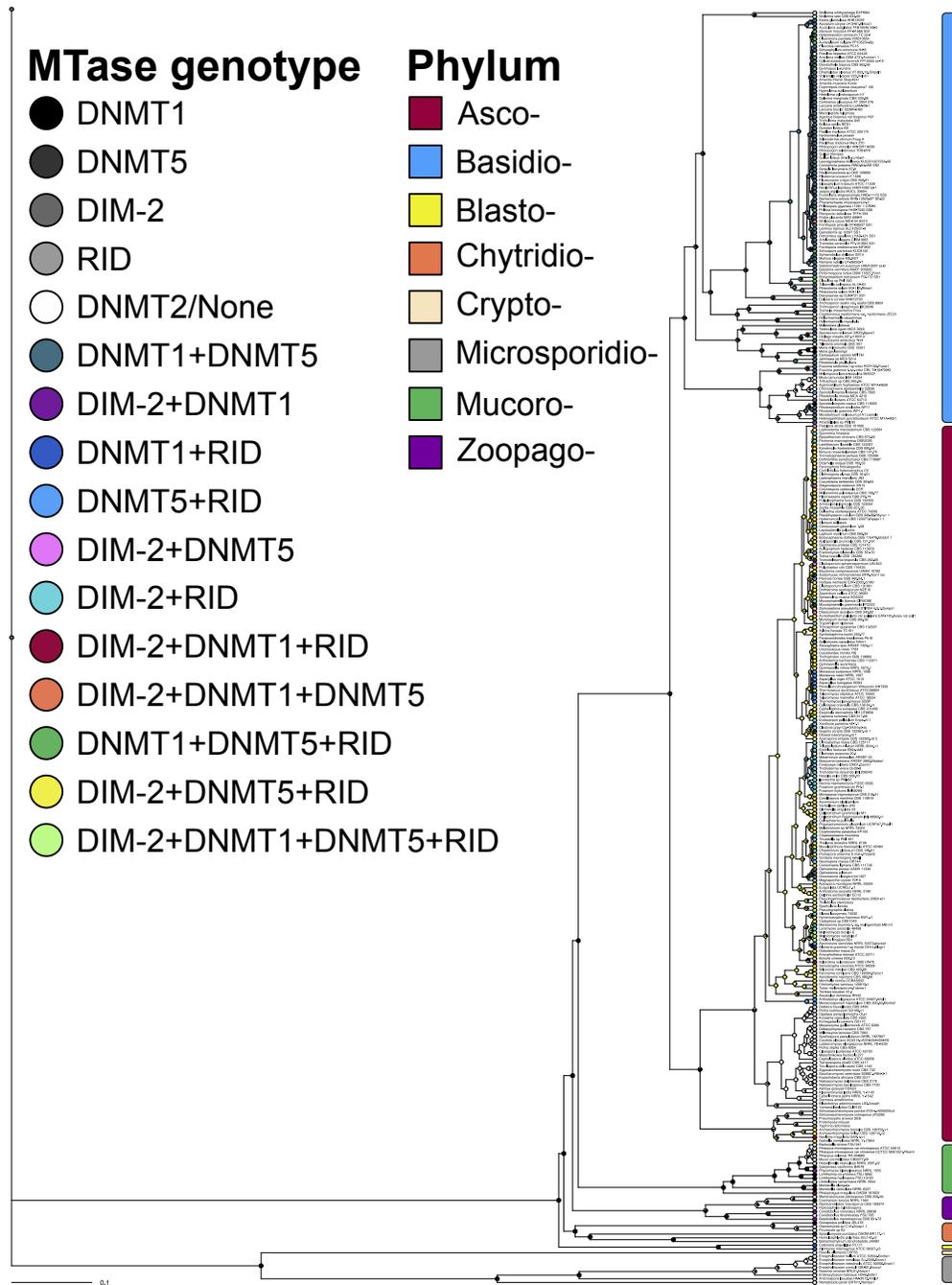
131

132

133

134

Supplementary Fig. 3



135

136

137

138

139

140 **Supplementary Figure 4**

141

142 **Individual and combinations of DNA MTases.**

143

144 The values within the areas of the Venn diagram indicate the number of fungal
145 species/strains with the corresponding DNA MTase genotype.

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

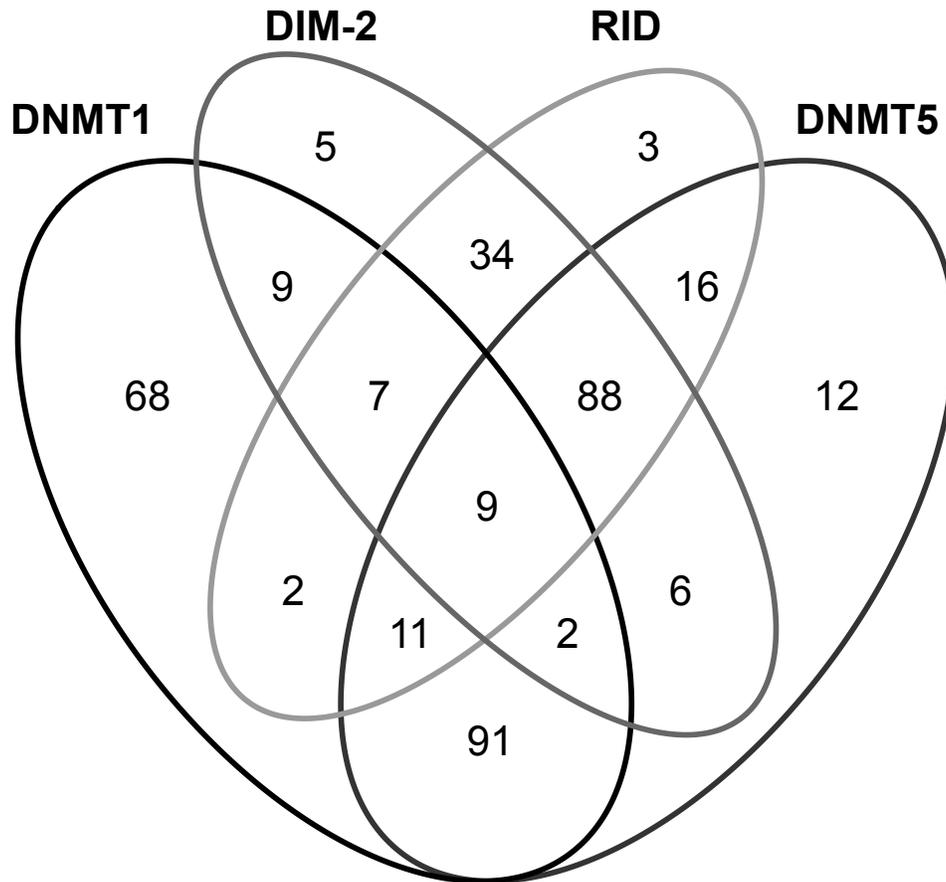
166

167

168

169

170



171
172
173
174
175

176 **Supplementary Figure 5**

177

178 **Predictors of methylation.**

179

180 Rank-order from best to worse predictors of genome-wide CG methylation.

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

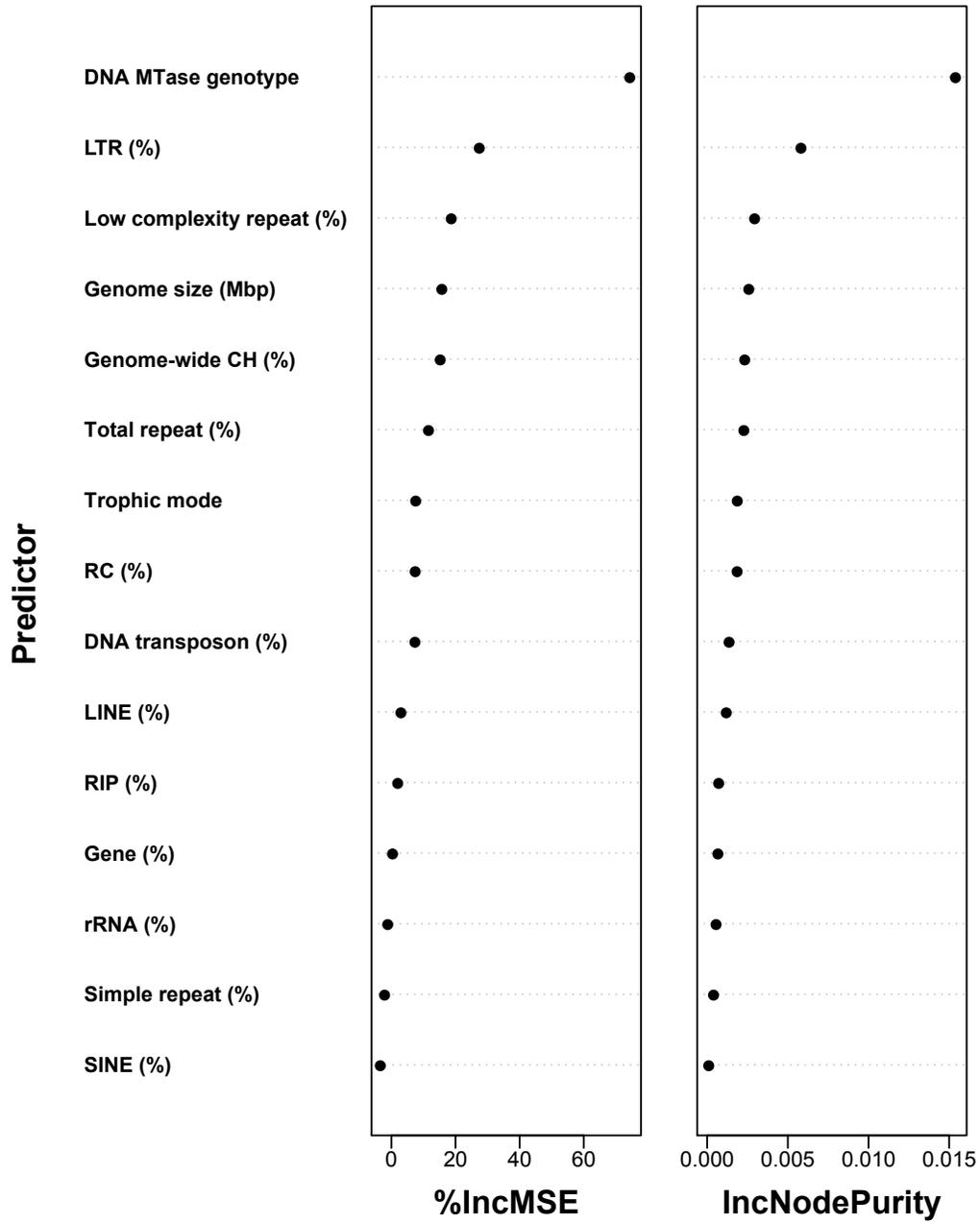
203

204

205

206

Supplementary Fig. 5



207

208

209

210

211

212 **Supplementary Figure 6**

213

214 **Explanations for discrepancies between DNA MTase genotype and**
215 **methylation levels.**

216

217 (a) Tissue-specific expression of DNA MTases in *Bot. cinerea* and *Pse.*
218 *destructans*. Grey-highlighted boxes specify tissues used for methylation
219 analyses. *a* and *b* above bars in *Bot. cinerea* specify RID in-paralogs. Error bars
220 correspond to the standard error of the mean (SEM). (b) Methylation in *Lep.*
221 *maculans* 'brassicae' is localized to short, gene-poor regions at the beginnings
222 and ends of supercontigs.

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

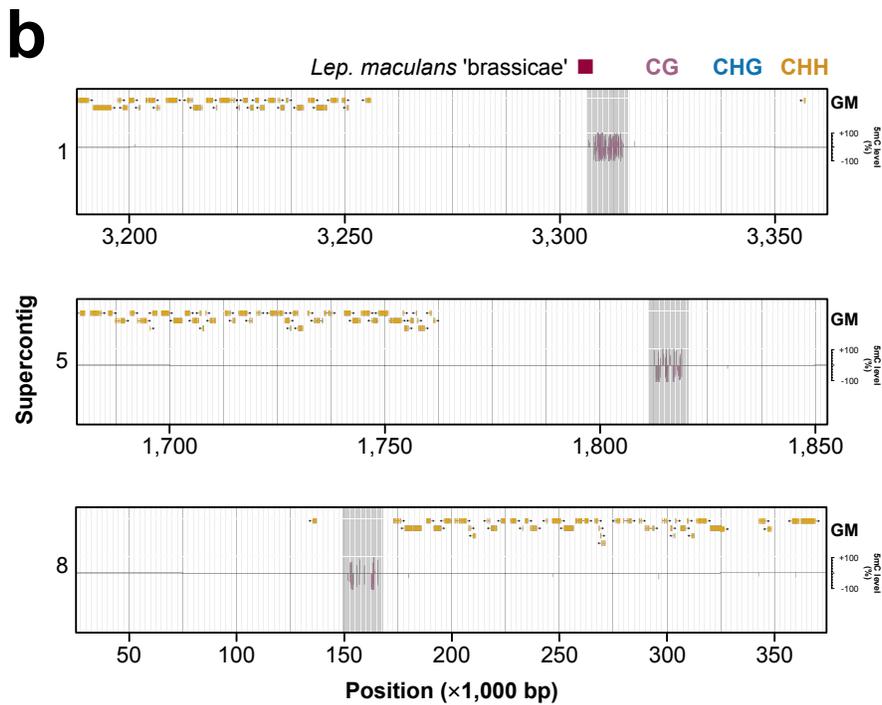
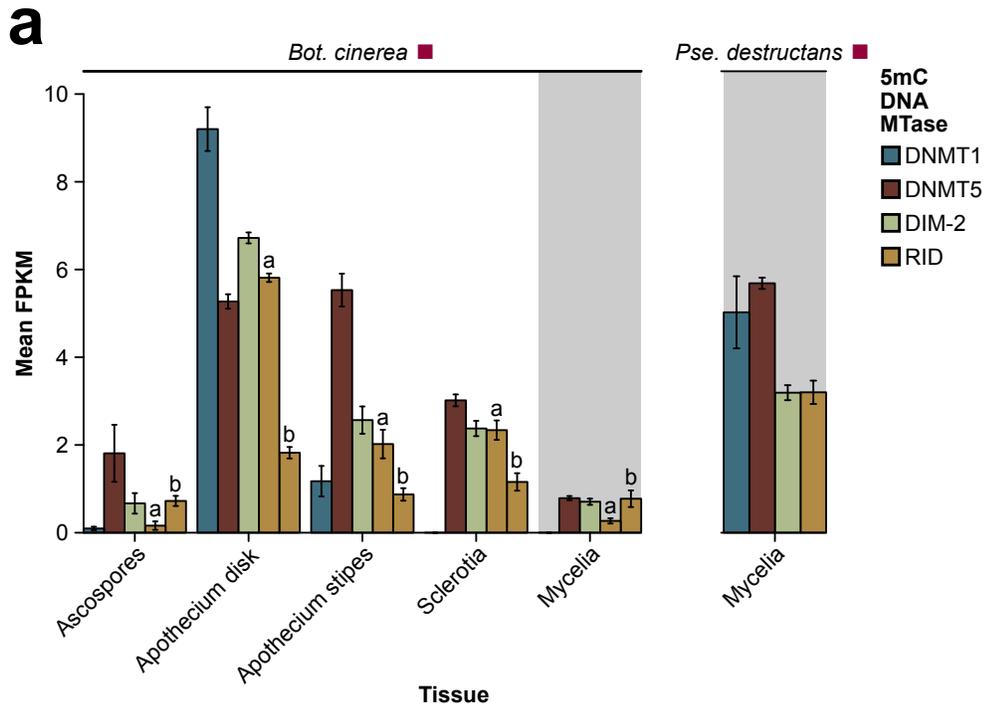
238

239

240

241

242



244

245

246

247

248

249 **Supplementary Figure 7**

250

251 **Context-specific methylation.**

252

253 Weighted methylation levels at all sequence contexts found across the genome
254 for fungal species not included in Fig. 2.

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

284

285 **Supplementary Figure 8**

286

287 **Genic and repeat and transposon methylation profiles.**

288

289 Weighted methylation at CG, CH, and CN sites upstream, within, and
290 downstream of all genes for fungal species not included in Fig. 4.

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

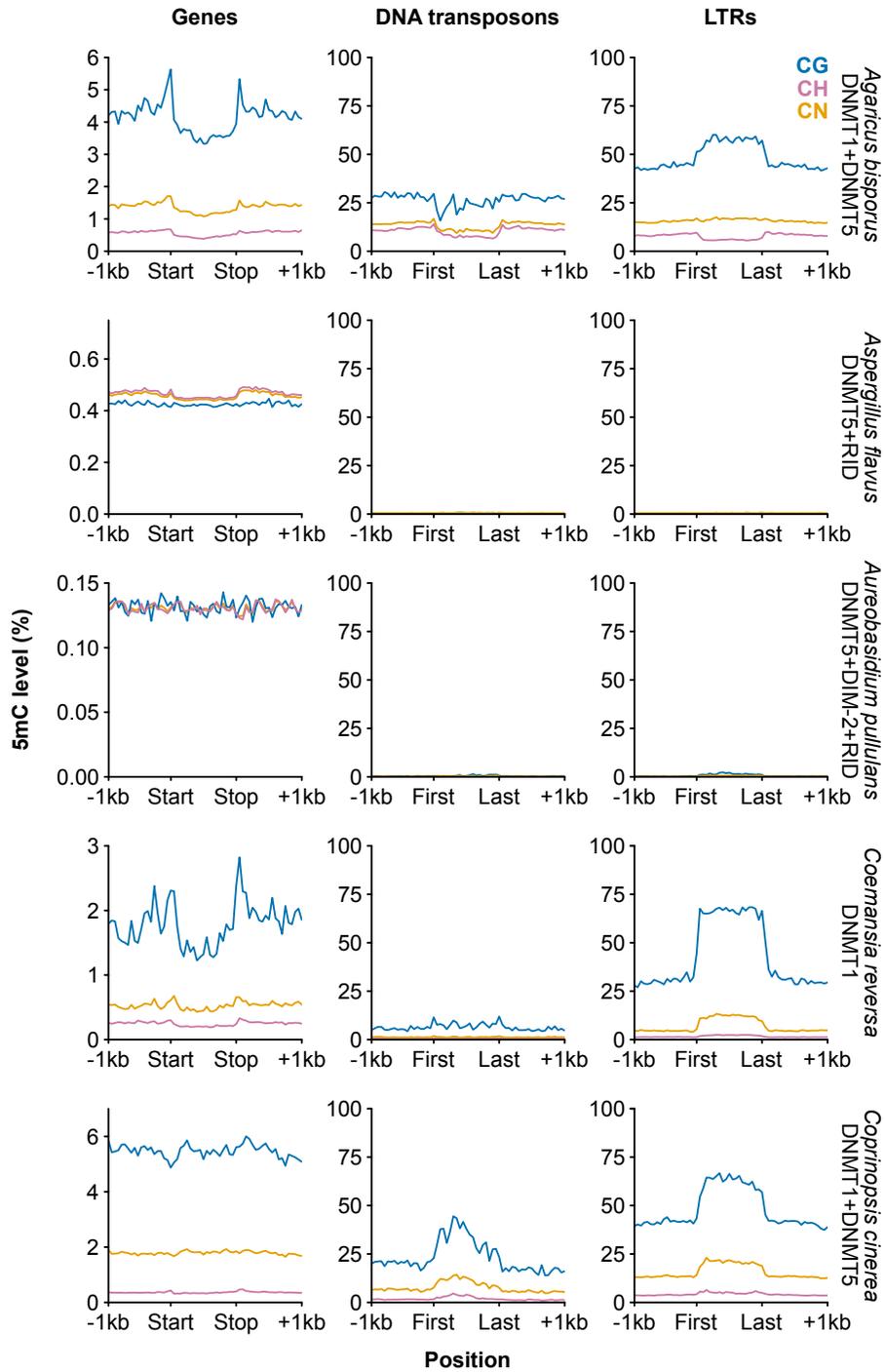
310

311

312

313

314



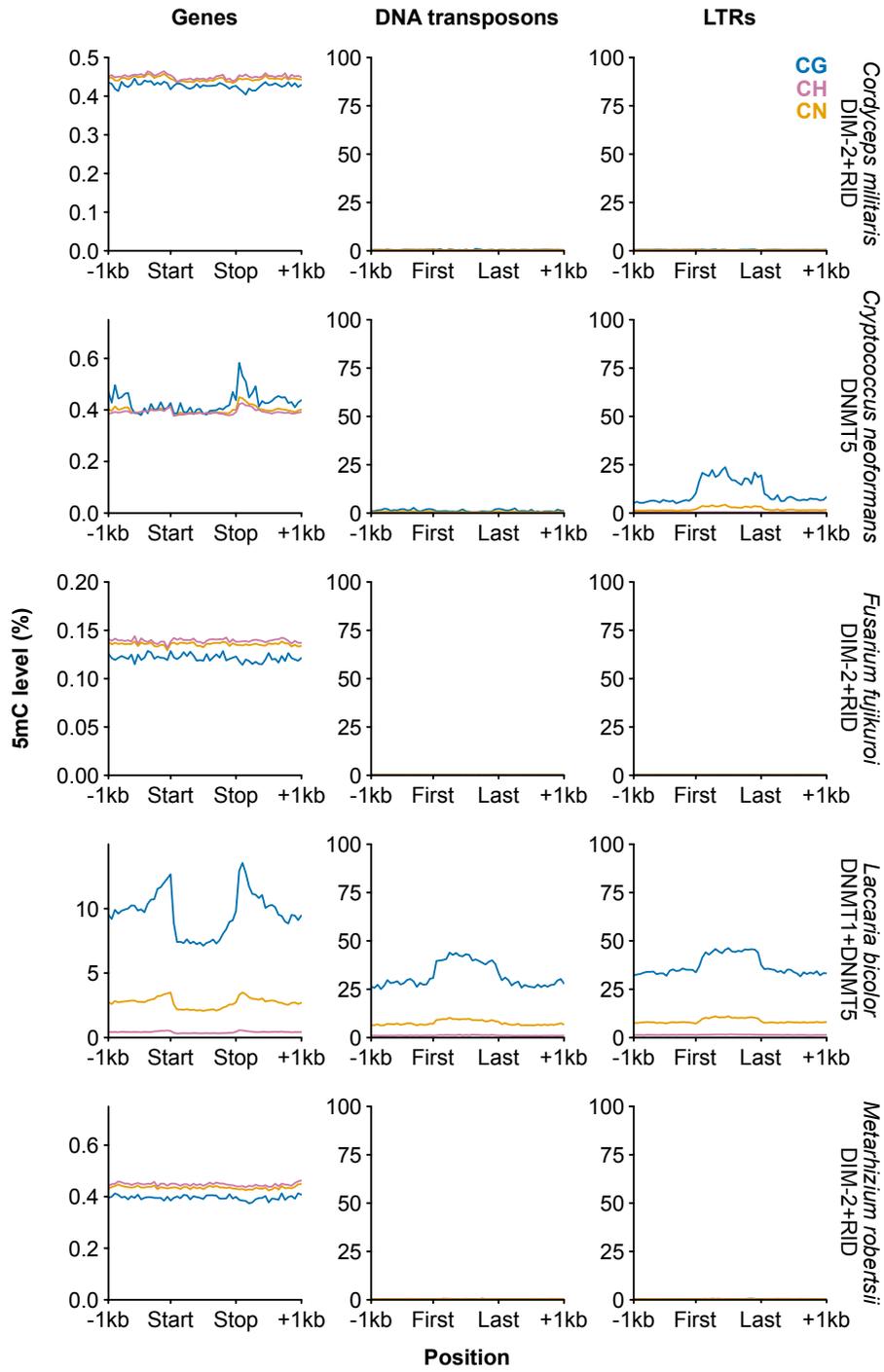
316

317

318

319

Supplementary Fig. 8



320

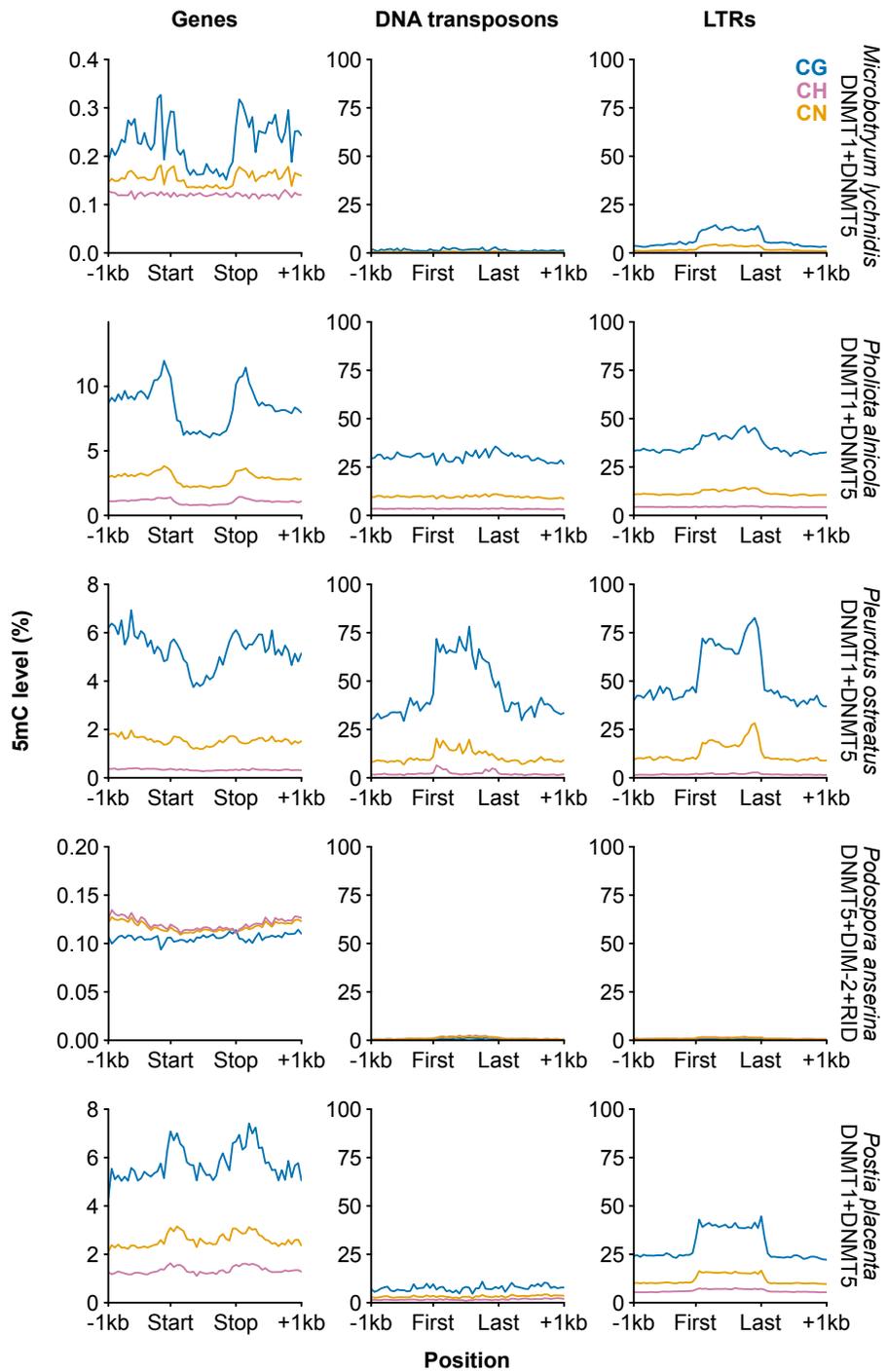
321

322

323

324

Supplementary Fig. 8



325

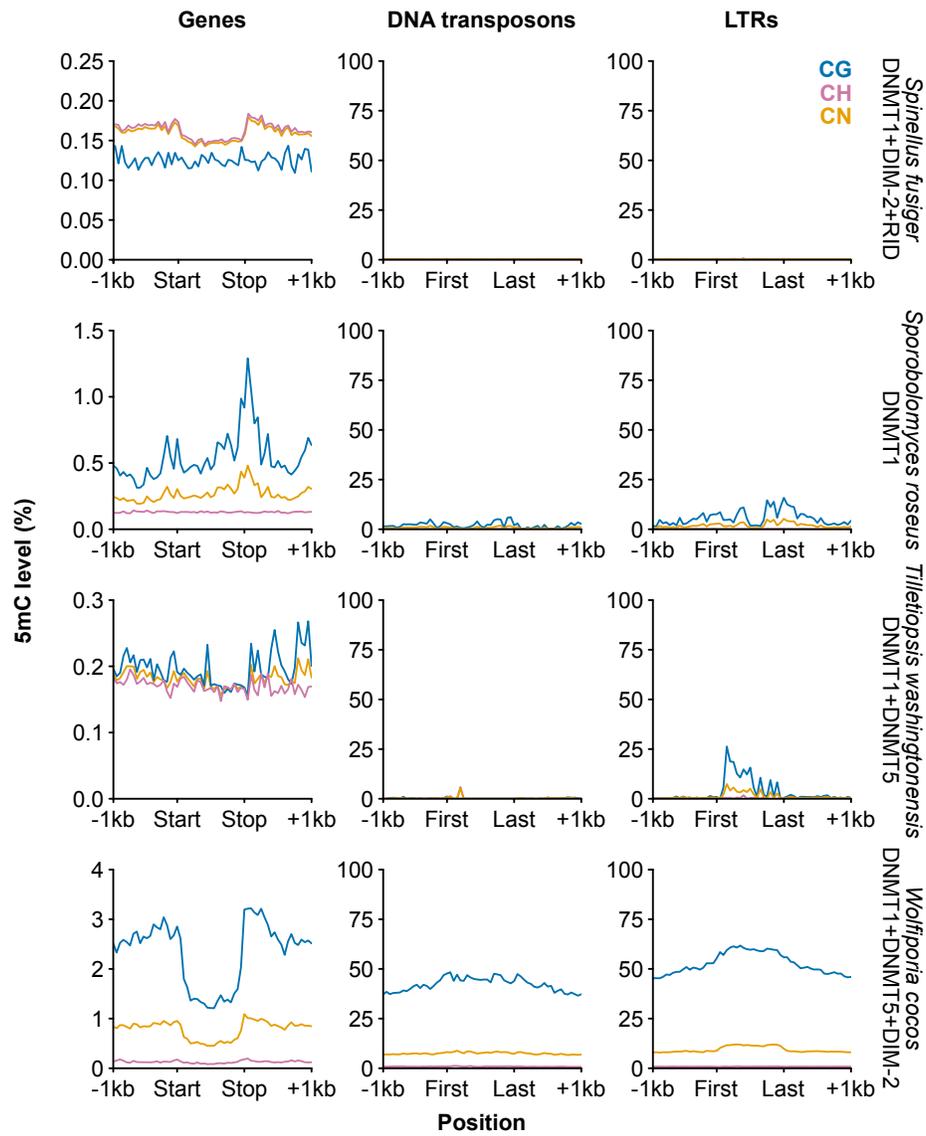
326

327

328

329

Supplementary Fig. 8



330

331

332

333

334

335 **Supplementary Figure 9**

336

337 **Genic 5mC density.**

338

339 Distribution of 5mC levels for genes for species not used in Fig. 3b.

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

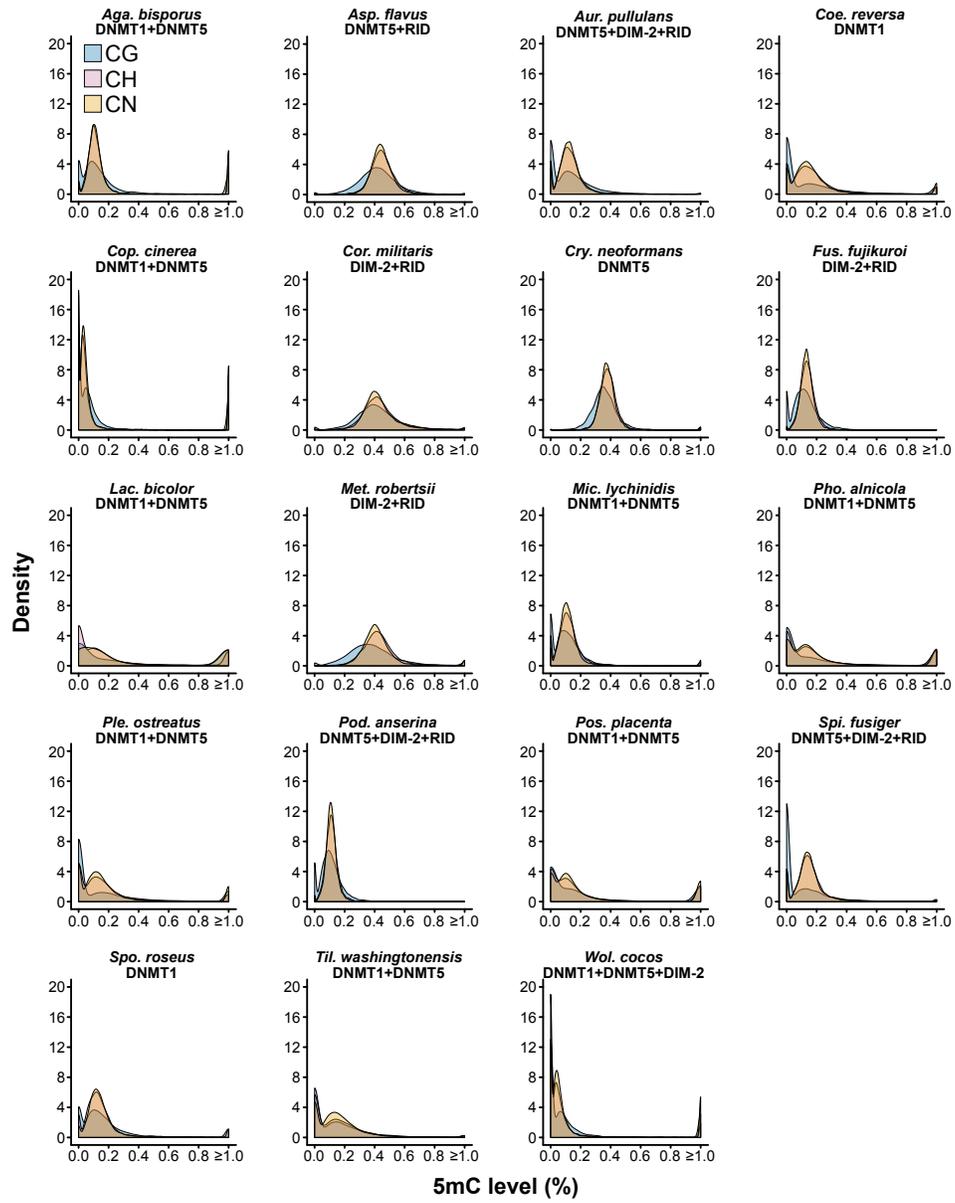
362

363

364

365

Supplementary Fig. 9



366

367

368

369

370

371 **Supplementary Figure 10**

372

373 **Relationship between gene expression (Fragments Per Kilobase of**
374 **transcript per Million mapped reads [FPKM]) and genic 5mC levels.**

375

376 Relationship between gene expression measured as FPKM and 5mC levels.

377 Empty deciles correspond to missing genic data for those 5mC levels. Boxplot

378 elements: center line, median; upper and lower "hinges", first and third quartiles

379 (the 25th and 75th percentiles), respectively; whiskers, 1.5× interquartile range.

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

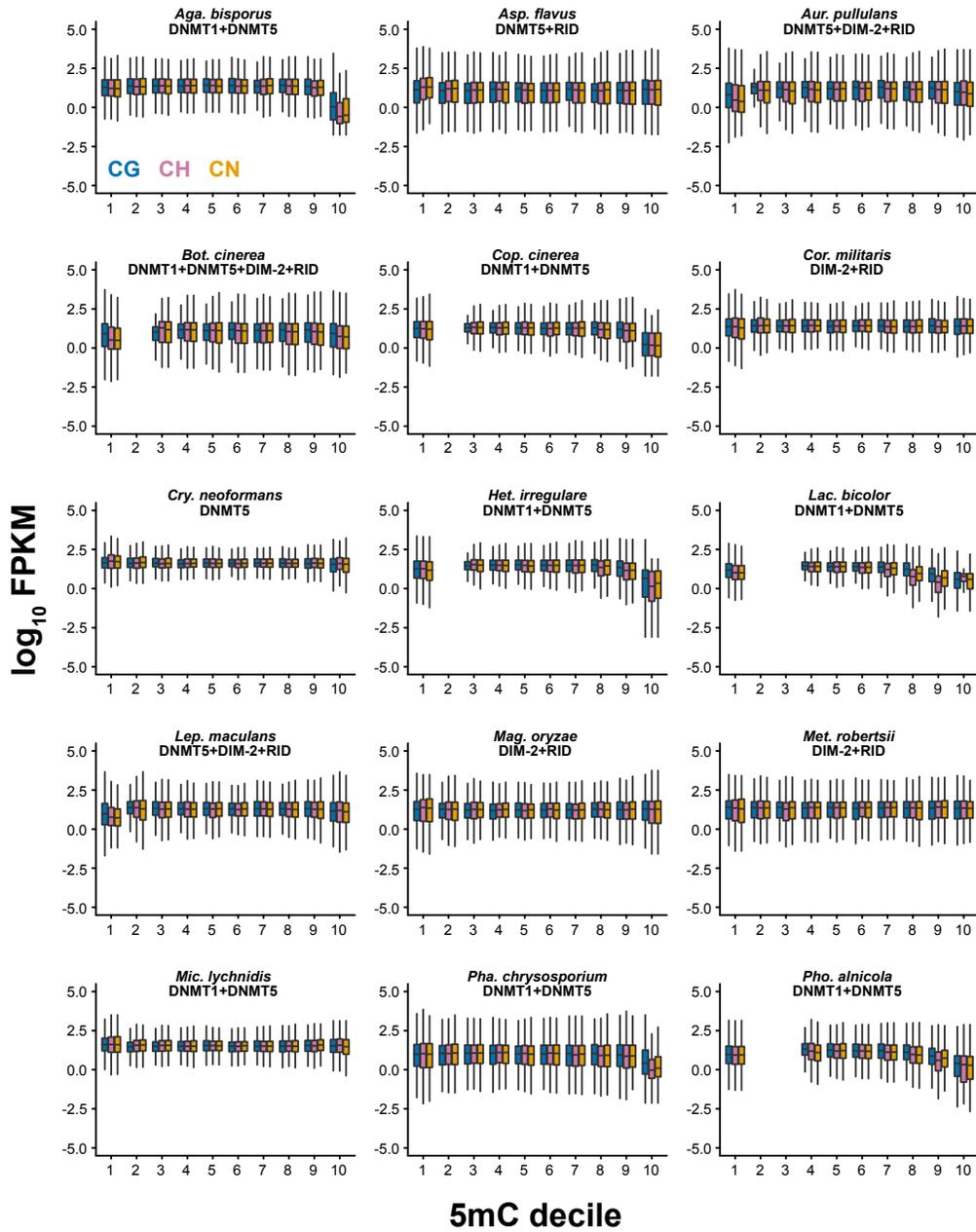
398

399

400

401

Supplementary Fig. 10

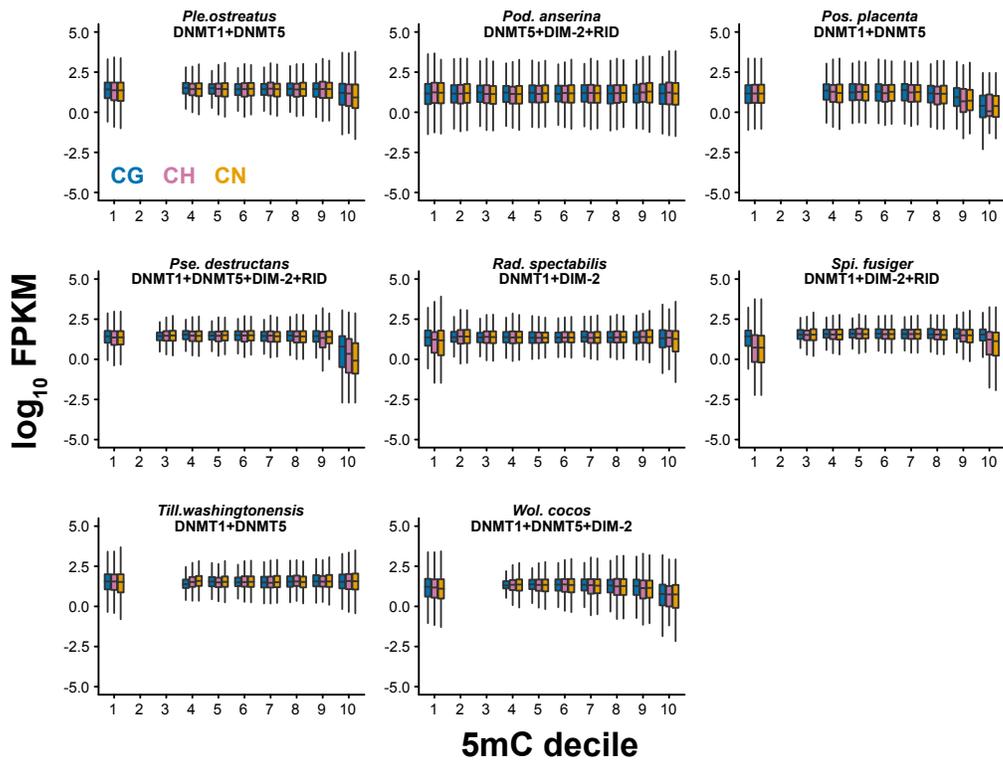


402

403

404

405



406

407

408

409

410

411 **Supplementary Figure 11**

412

413 **DNA transposon and Long Terminal Repeat (LTR) 5mC density.**

414

415 Distribution of 5mC levels for DNA transposons and LTRs for species not used in
416 Fig. 4b and c.

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

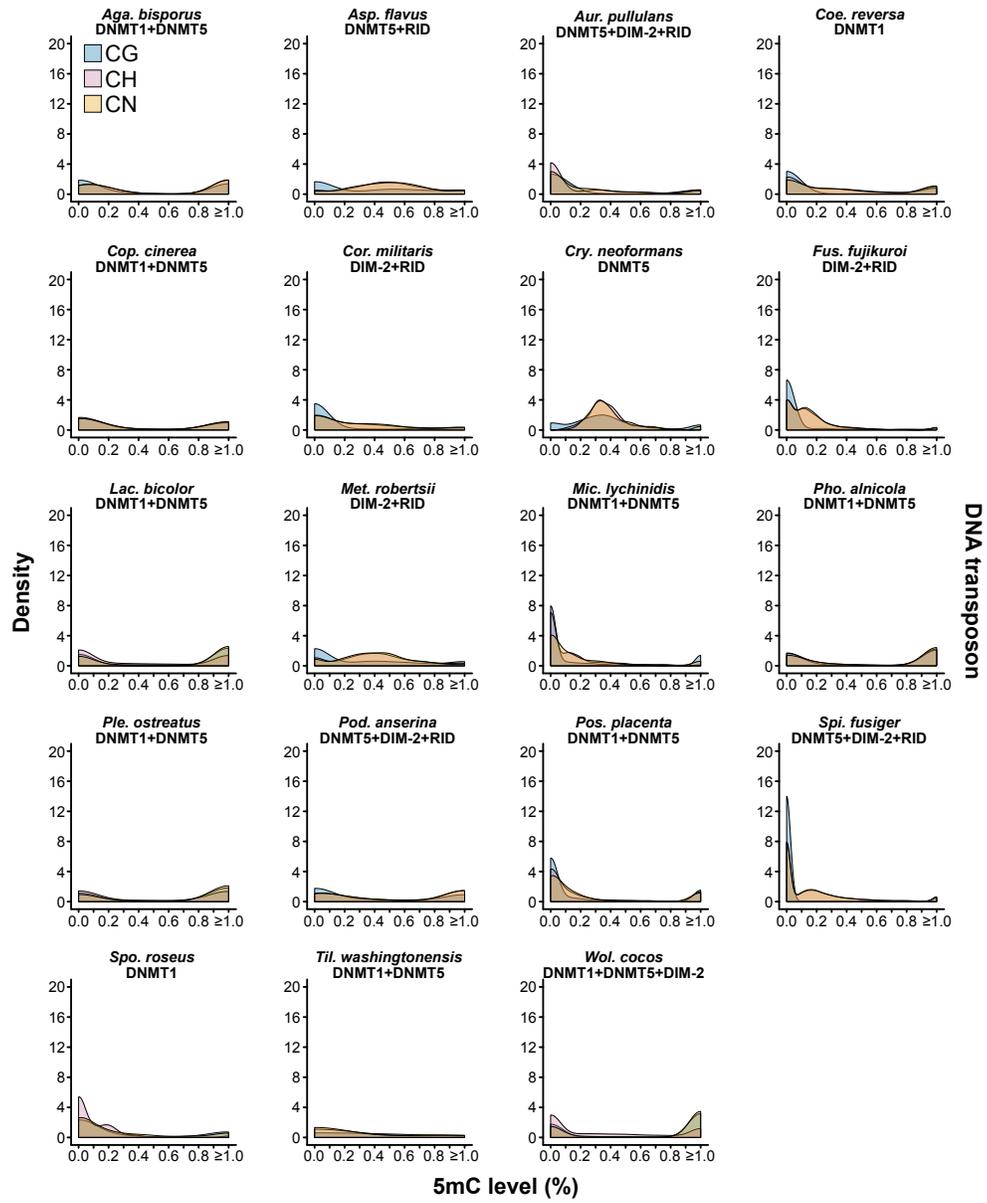
438

439

440

441

Supplementary Fig. 11



442

443

444

445

446

447 **Supplementary Figure 12**

448

449 **Methylated Cytosine Clusters.**

450

451 Methylation level across all contexts and length distributions of Methylated
452 Cytosine Clusters (MCCs), and proportion of genome occupied by methylated
453 clusters for 17 fungal species. Boxplot elements: center line, median; upper and
454 lower "hinges", first and third quartiles (the 25th and 75th percentiles),
455 respectively; whiskers, 1.5× interquartile range; large points, outliers.

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

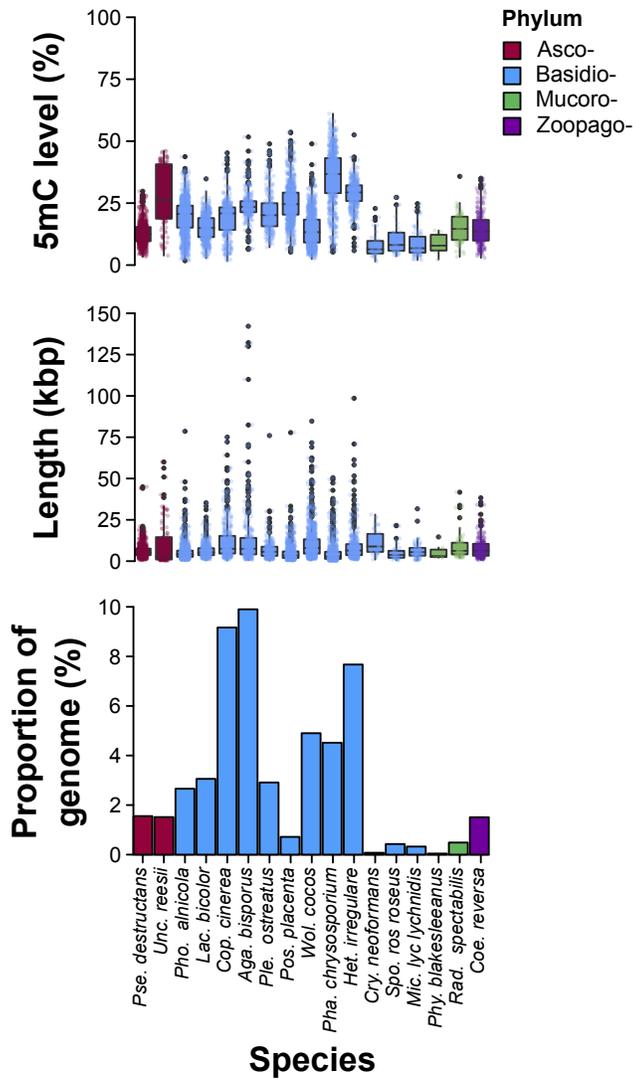
473

474

475

476

477



478

479

480

481

482

483 **Supplementary Figure 13**

484

485 **Gene Ontology (GO) term enrichment of genes located in MCCs.**

486

487 Proportion of statistically significant (P value <0.05) GO terms for genes located
488 in MCCs.

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

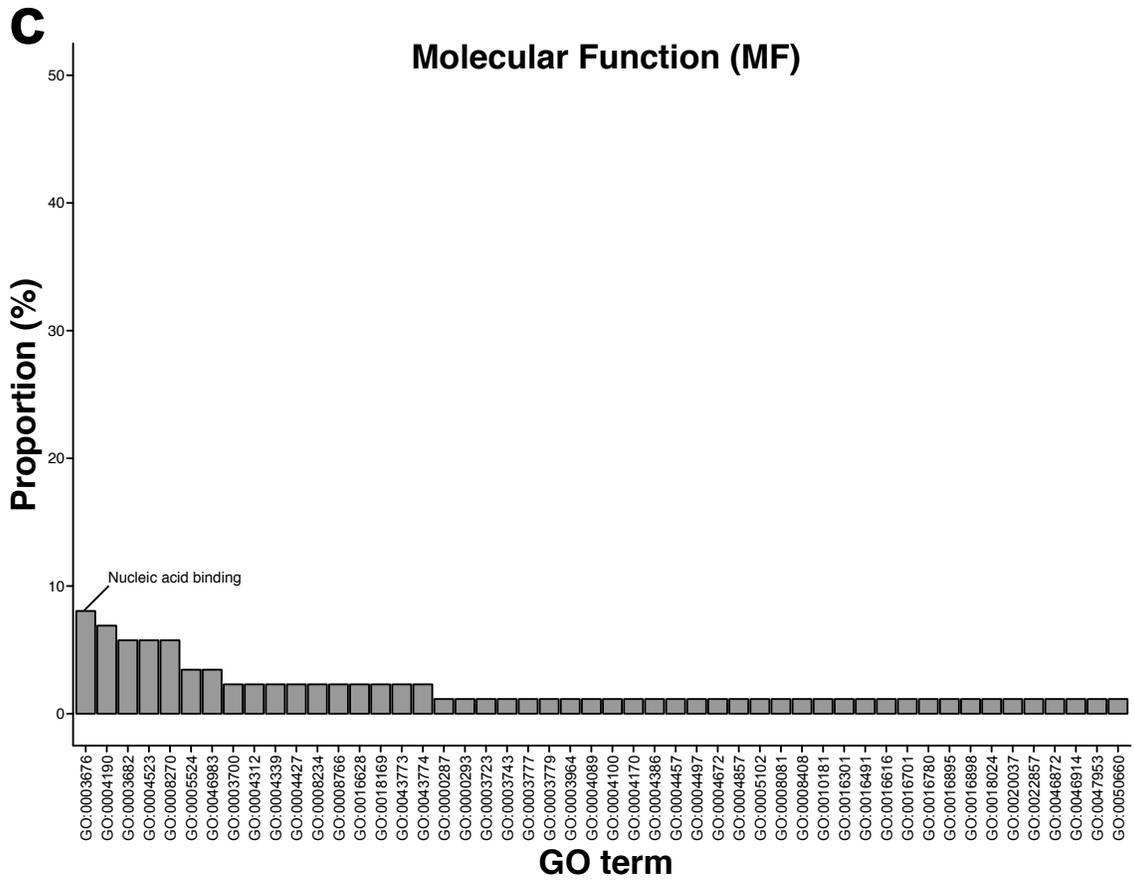
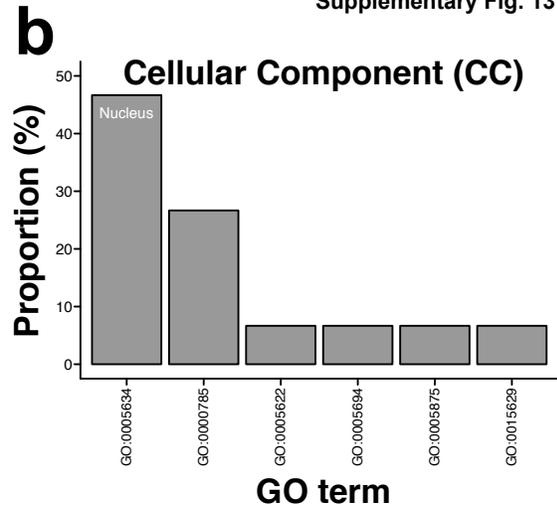
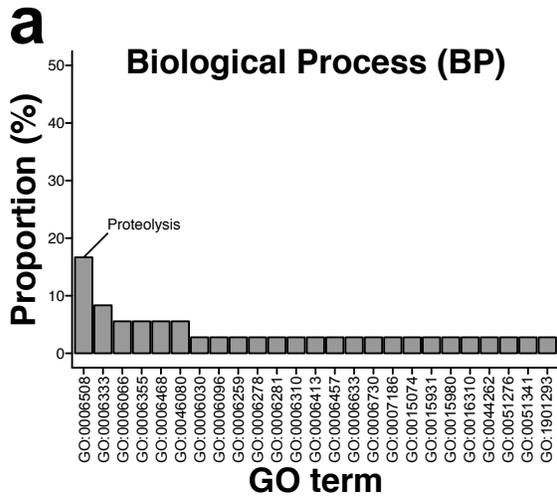
509

510

511

512

513



514
515
516
517
518

519 **Supplementary Figure 14**

520

521 **Repeat and transposon content.**

522

523 The genomic proportion of repeats and transposons for 40 fungal species.
524 Repeats were detected and annotated using REPET v2.5. RIP loci were detected
525 and annotated using the RIP product index, the RIP substrate index, and
526 composite RIP index. See **ONLINE METHODS** for further details on repeat and
527 transposon detection and annotation.

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

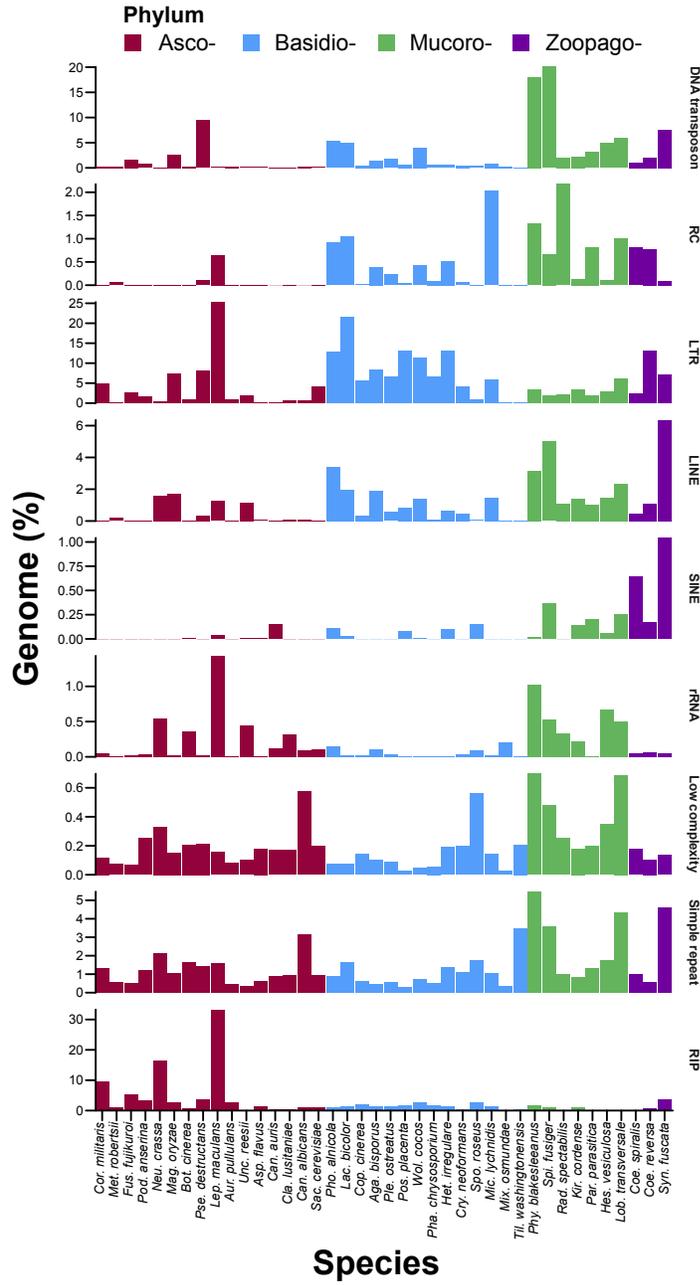
545

546

547

548

549



550
 551
 552
 553
 554

555 **Supplementary Figure 15**

556

557 **ALKBH phylogeny.**

558

559 Phylogenetic relationships and ALKBHs of Chordata, fungi, and Nematoda.

560 Values at selected nodes indicate posterior probability. Branch lengths are in

561 units of amino acid substitutions per amino acid site. Unlabeled clades represent

562 taxa-specific clades of ALKBHs.

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

591 **Supplementary Figure 16**

592

593 **METTL phylogeny.**

594

595 Phylogenetic relationships and METTLs of fungi. Values at selected nodes
596 indicate posterior probability. Branch lengths are in units of amino acid
597 substitutions per amino acid site.

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

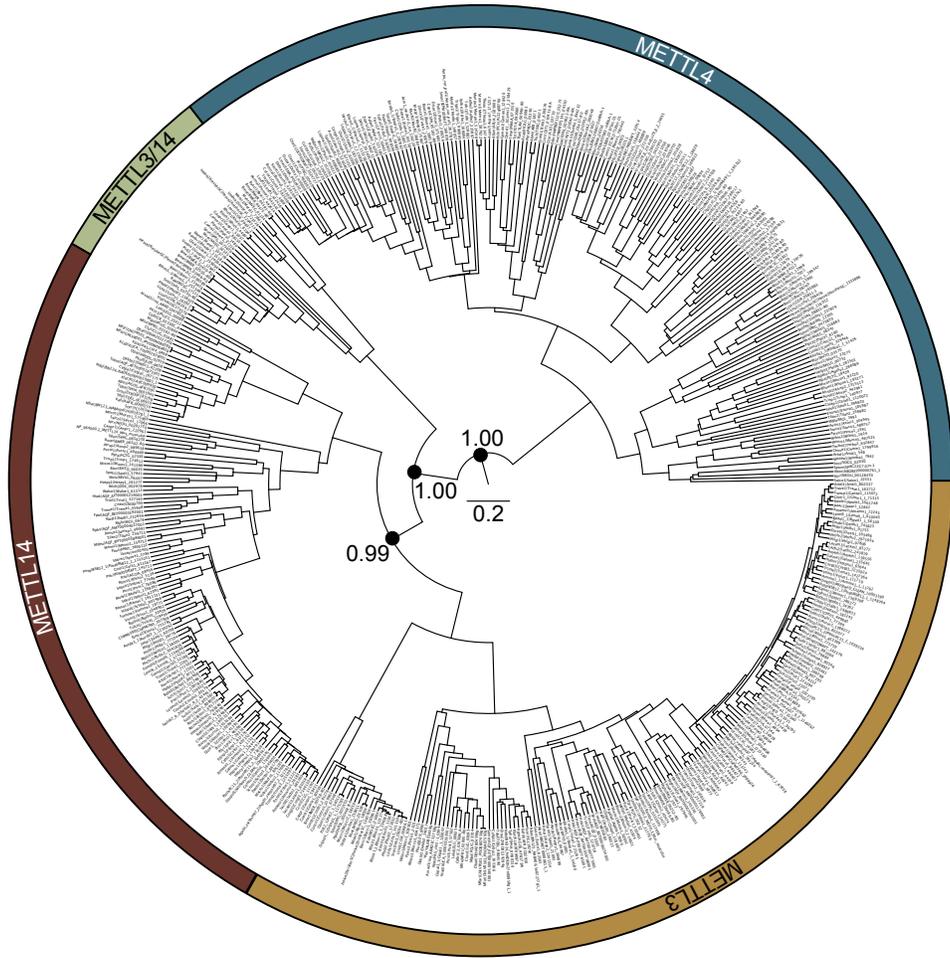
617

618

619

620

621



MTase

- METTL3
- METTL14
- METTL3/14
- METTL4

622

623

624

625

626

627 **Supplementary Table 1 | 5mC DNA MTase and the tRNA methyltransferase**
628 **DNMT2 annotations for 528 fungal species investigated in this study.**

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658 **Supplementary Table 2 | WGBS and mapping statistics for 40 fungal**
659 **species investigated in this study.**

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689 **Supplementary Table 3 | Number of species per phylum for each observed**
690 **5mC MTase genotype.**

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719 **Supplementary Table 4 | 5mC DNA MTase and the tRNA methyltransferase**
720 **DNMT2 annotations for a subset of animalia, chlorophyta, fungi, and**
721 **prokaryota. Protein models correspond to those used in Supplementary**
722 **Fig. 2.**

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749 **Supplementary Table 5 | Number of CG-, CH, and CN-enriched genes**
750 **across fungal species investigated.**

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779 **Supplementary Table 6 | ALKBH annotations for Chordata, fungi, and**
780 **Nematoda investigated in this study. Protein models correspond to those**
781 **used in Supplementary Fig. 15.**

782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809

810 **Supplementary Table 7 | Results from Pagel's test for correlated evolution.**

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841 **Supplementary Table 8 | Results from Phylogenetic Generalized Least**
842 **Squares (PGLS).**

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870 **Supplementary Table 9 | Annotated proteins from fungal species**
871 **containing the N-6 DNA Methylase domain (PF02384) as identified by**
872 **Interproscan v5.23-62.0 (45).**

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901 **Supplementary Table 10 | METTL annotations for fungal species**
902 **investigated in this study. Protein models correspond to those used in**
903 **Supplementary Fig. 16.**

904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929

930 **Supplementary Table 11 | Annotated proteins for fungal species containing**
931 **the domain the methyltransferase small domain (i.e., N6AMT1 proteins) as**
932 **identified by Interproscan v5.23-62.0 (45).**